

A Critical Examination of Summary Measures of Population Health

Christopher JL Murray

Joshua A Salomon

Colin Mathers

Summary measures of population health are measures that combine information on mortality and non-fatal health outcomes to represent the health of a particular population as a single number (Field and Gold 1998). Efforts to develop summary measures of population health have a long history (Sanders 1964, Chiang 1965, Sullivan 1966, 1971, Piot and Sundaresan 1967, Fanshel and Bush 1970, Katz et al. 1973, Ghana Health Assessment Project Team 1981, Preston 1993). In the past decade, there has been a markedly increased interest in the development, calculation and use of summary measures. The volume of work from members of the Réseau de Espérance de Vie en Santé (REVES) offers one indication of the activity in this field. (Robine et al. 1993, Mathers et al. 1994). Applications of measures such as active life expectancy (Katz et al. 1983) have been numerous, especially in the United States. Calculations of related summary measures such as Disability-Free Life Expectancy (DFLE) also have appeared frequently in recent years (Bone 1992, Bronnum-Hansen 1998, Crimmins et al. 1997, Mutafova 1997, Sihvonen et al. 1998, Valkonen et al. 1997). Another type of summary measure, Disability-Adjusted Life Years (DALYs) has been used in the Global Burden of Disease Study (Murray and Lopez 1996a, 1996b, 1996c, 1997a, 1997b, 1997c, 1997d) and in a number of National Burden of Disease Studies (e.g. Lozano et al 1994, 1995, Fundacion Mexico de la Salud 1995, Republica de Colombia Ministerio de Salud 1994, Ruwaard and Kramers 1998, Bowie et al 1997, Concha et al 1996, Murray et al. 1997, Murray et al. 1998). Reflecting this rising interest in the academic and policy communities, the United States' Institute of Medicine convened a panel on summary measures and published a report that included recommendations to enhance public discussion of the ethical assumptions and value judgements, establish standards, and invest in education and training to promote use of summary measures. (Field and Gold 1998).

Interest in summary measures relates to a range of potential applications of them. At least eight uses are worth highlighting here:

1) *Comparing the health of one population to the health of another population*

Such comparative judgements are essential to evaluations of the performance of different health systems. Comparisons may allow decision-makers to focus their attention on those health systems with the worst performance. In addition, comparative judgments provide the possibility of analyzing the key contributors to differences in health between populations.

2) *Comparing the health of the same population at different points in time.*

Monitoring changes in health status over time is essential for the evaluation of health system performance and progress towards stated goals for a given society.

3) *Identifying and quantifying overall health inequalities within populations.*

4) *Providing appropriate and balanced attention to the effects of non-fatal health outcomes on overall population health.*

In the absence of summary measures, conditions that cause decrements in function but not mortality tend to be neglected relative to conditions that primarily cause mortality.

5) *Informing debates on priorities for health service delivery and planning.*

When a summary measure is combined with information on the contributions of different causes of disease and injury or risk factors to the total, such information

should be a critical input to debates on the identification of a short-list of national health priorities that will consume the attention of senior managers in public health agencies and government leaders.

6) *Informing debates on priorities for research and development in the health sector.*

The relative contributions of different diseases, injuries and risk factors to the total summary measure is also a major input to debate on priorities for research and development investment (World Health Organization 1996).

7) *Improving professional training curricula in public health.*

8) *Analyzing the benefits of health interventions for use in cost-effectiveness analyses.*

The change in some summary measure of population health offers a natural unit for quantifying intervention benefits in these analyses.

Broad interest and use of summary measures in the policy demonstrates the recognition of their value at the practical level for many of the eight purposes identified here (repeat NBD citations, Brundtland 1998, World Bank 1993, Bobadilla 1996).

Despite the many uses of summary measures, Williams (1999) and Mooney (1998) have argued that summary measures of population health are not informative – for a more detailed response to Williams see Murray and Lopez (1999). According to this argument, only information on the incremental costs and benefits of interventions, and not information on the level of health in a population, should be considered relevant inputs to decision-making. The merit of this argument is perhaps best evaluated by considering analogous arguments to employment or national accounts. Applying the logic of Williams and Mooney, we might conclude that the unemployment rate has no policy relevance, and that only the incremental costs and the number of new jobs created through different government policy options would be informative. Similarly, Jamison (1996) has drawn an analogy to National Income and Product Accounts, where again, a marginalist view would suggest that these measurements are not worthwhile, and that only the costs and benefits of all possible economic policy options should be collected. In general, Williams' and Mooney's line of reasoning would find few if any proponents among economists, and would seem to be equally unconvincing in the realm of health policy.

Before proceeding with discussions on the types of summary measures and the question of choosing among them, it is important to distinguish clearly between the nature and quality of various inputs to summary measures and the properties of summary measures themselves. Information on age-specific mortality and the epidemiology of non-fatal health outcomes provides a basic input to any type of summary measure. Another critical input is information on the values attached to various health states relative to ideal health or death. Health status measurement instruments such as SF-36 can be used to describe health states in terms of performance on various domains of health. This information could be combined with health state valuations to calculate the non-fatal health component of many different summary measures, but SF-36 and other health status measurement instruments do not themselves constitute summary measures of population health. As noted below, the occasional confusion surrounding the distinction between data inputs to a summary measure and the summary measure itself may be exacerbated when

summary measures are linked **by definition** to particular health status instruments, as in Years of Healthy Life (Erickson et al. 1995). A more fundamental problem is that comparisons between communities or cultural groups within communities are extremely difficult to interpret using the results of most existing health status instruments. The challenge of understanding and accounting for the difference between self-assessed and observed measures of health status is not the focus of this paper but is a critical empirical challenge to the enterprise of measuring population health.

In examining the properties of various summary measures, it is important to bear in mind the ultimate goal of influencing the policy process (Field and Gold 1998). Because of their potential influence on international and national resource allocation decisions, summary measures must be considered as normative measures. As stated by the IOM panel, “all measures of population health involve choices and value judgements in both their construction and their application.” Great care must be taken in the construction of summary measures precisely because they may have far-reaching effects. Normative aspects of the design of summary measures have been the subject of extensive debate (see, for instance, Brock 1998, Murray 1996). As one example, questions of summary measures and concerns for equity have been given considerable attention (Anand and Hanson 1998, Daniels 1998, Williams 1997, 1999).

In the remaining sections of this paper, we begin by presenting a typology of summary measures in terms of two broad families: health expectancies and health gaps. We then outline key attributes by which various members of these families may be distinguished. We propose a number of criteria and other properties which can be used to evaluate different summary measures. Finally, we discuss some of the broad implications of this framework for choosing summary measures and consider directions for future progress.

A typology of summary measures of health

A wide array of summary measures have been proposed (for example, active life expectancy, disability-free life expectancy, dementia-free life expectancy, disability-adjusted life expectancy, health-adjusted life expectancy, healthy life-years, Years of Healthy Life, disability-adjusted life years, *etc.*). On the basis of a simple survivorship curve, these measures can be divided broadly into two families: health expectancies and health gaps. The bold curve in Figure 1 is an example of a survivorship curve for a hypothetical population. The survivorship curve indicates, for each age along the x-axis, the proportion of an initial birth cohort that will remain alive at that age.

The area under the bold survivorship curve ($A+B$) represents life expectancy at birth. Health expectancies are measures of this area that take into account some lower weights for years lived in health states worse than full health, represented as the area labeled B in the diagram. More formally,

$$\text{Health expectancy} = A + f(B)$$

where $f(\cdot)$ is some function that assigns weights to health states less than ideal health using a scale on which full health has a weight of 1.

In contrast to health expectancies, health gaps quantify the *difference* between the actual health of a population and some stated norm or goal for population health. The health goal implied by Figure 1 is for everyone in the entire population to live in ideal

health until the age indicated by the vertical line enclosing area C at the right.ⁱ The health gap, therefore, is a function of area C and area B, or more formally,

$$\text{Health gap} = C + g(B)$$

where $g(\cdot)$ is some function that assigns weights to health states less than full health, using a scale on which a weight of 1 implies that time lived in a particular health state is equivalent to time lost due to premature mortality. In the literature, many different health expectancies and health gaps have been proposed and calculated for selected populations.

Absolute and Covariate Independent Forms of Summary Measures

While the survivorship function in Figure 1 is a convenient heuristic to illustrate the difference between health expectancies and health gaps, it is important to recognize that summary measures may take either an absolute or an age-independent form. By their construction, health expectancies are measures that do not depend on the particular age structure of a population. Health gaps, on the other hand, are usually expressed in absolute terms and as such are age-structure-dependent. For example, a health gap may be expressed as the total number of healthy years of life that have been lost in a population, which will vary according to the particular age distribution of the members of that population. Familiar methods such as age-standardization may be used, however, to compute health gaps that are independent of population age structure. It is also possible to conceive of health expectancies that *do* depend on age-structure – for example, total healthy years lived in a population – but these constructs do not seem as intuitively appealing, and they evoke other contentious issues such as the use of period vs. cohort measures.

Most discussions of health statistics have focused on the development of age-structure independent measures, but age obviously is only one of innumerable covariates of health outcomes. We might therefore imagine developing sex-independent, race-independent or income-independent forms of summary measures just as we have age-structure-independent forms. There are several arguments, however, for paying special attention to age-structure independent forms of summary measures. First, age is perhaps one of the most powerful determinants of health outcomes, such that comparisons based on measures that are not age-structure independent may be dominated, in some cases, by variation in this variable. Second, we cannot realistically intervene on age; further, age is unique in that every individual has the possibility of belonging to each age. Although we could imagine applying at least some of these arguments to other factors such as sex and race, it is likely that there will always be an interest in summary measures that are independent of age-structure. The design of a summary measure and the range of covariate-independent forms of the summary measure that might be developed will ultimately depend critically on the intended use for the summary measure.

Attributes of Health Expectancies

A wide range of health expectancies have been proposed since the original notion was developed (Sanders 1964), including active life expectancy (Katz et al.1983), disability-free life expectancy (Robine et al. 1993, Mathers et al. 1994), disability-adjusted life expectancy (Murray and Lopez 1997), years of healthy life (Erickson et al. 1995), quality-adjusted life expectancy (Fanshel and Bush 1970, Wilkins and Adams 1992) and dementia-free life expectancy (Ritchie et al. 1993). Although not proposed as a summary measure, Cutler and Richardson (1997,1998) have developed

a type of health expectancy called health capital. These various health expectancies can be distinguished on the basis of five aspects.

First, as with standard life tables (Shyrock and Siegel 1971), health expectancies can be calculated for a period or for a cohort. The more common period method calculates the health expectancy for a hypothetical birth cohort exposed to currently observed event rates (e.g., rates of mortality, incidence and remission) over the course of their lifetime. As long time-series datasets on the epidemiology of non-fatal health outcomes are not available, we are not aware of any calculations of cohort health expectancies for real populations. Deeg and colleagues (1994) projected disability transition rates based on longitudinal data for the Netherlands but have not converted these into cohort health expectancies. Barendregt and Bonneux (1994) have calculated changes in cohort disease-free life expectancy due to hypothetical interventions in a simulation model.

Second, health expectancies may be distinguished based on the calculation method used. The pioneering efforts by Sullivan (1966) and others to estimate health expectancies used applications of the prevalence-rate life table borrowed from working life, marriage and education life tables (e.g., Wolfbein 1949). Katz et al. (1983) proposed that calculations of active life expectancy be based on double-decrement life tables, in which individuals can move into two absorbing states, limited function and death. More recently, multi-state life tables have been estimated for health expectancies (Branch et al. 1991, Rogers et al. 1990, Mathers and Robine 1997a). Robine et al. (1992) and Barendregt et al. (1994) have argued that the multi-state method is required logically so that health expectancy is based only on currently measured mortality, incidence and remission and not on prevalence. Robine et al. (1992) argue that prevalence is not a period measure; it is a stock variable rather than a flow. The strength of this argument is addressed below. In real populations, there may not be much differences between health expectancies calculated using the prevalence, double decrement or multi-state approaches (Mathers and Robine 1997).

Third, perhaps the most important variation across health expectancies is in the definition and measurement of health status. Many health expectancies are linked to a particular health status measurement instrument. Active life expectancy is linked to variants of the activities of daily living. The Years of Healthy Life measure is linked to two questions collected on the U.S. National Health Interview Survey (NHIS), one on activity limitations and one on perceived general health (Erickson et al. 1995, Erickson 1998). And the quality-adjusted life expectancy of Wilkins and Adams (1992) is linked to a question on activity restriction on the Canada Health Survey. In other cases such as dementia-free life expectancy, the health expectancy is linked to a particular diagnosis or a single domain of health. Disability-free life expectancy is often calculated from data on long-term disability and includes the duration of a condition in its definition of disability. NHIS data on self-assessed general health has been used in the calculation of health capital (Cutler and Richardson 1997), although this measure is not, by definition, linked strictly to this instrument. Clearly, wherever a health expectancy is defined in reference to a particular instrument, then the summary measure will depend critically on the reliability and validity of that instrument. All of the particular instruments mentioned here represent very limited conceptions of health, emphasizing a restricted set of physical domains, in contrast to other health status instruments more widely applied in current practice, such as EuroQol or SF-36, which capture multiple dimensions of health. Direct linkages are not required in the

construction of summary measures, so they may complicate evaluation of the properties of summary measures unnecessarily.

Fourth, health expectancies also may be distinguished by the method used to assign values to time spent in health states worse than ideal health. At one level, some measures such as ALE and DFLE use dichotomous valuations as illustrated in Figure 2. Up to some arbitrary threshold the valuation is zero (i.e., equivalent to the valuation of death), and beyond this threshold the valuation is one (i.e., equivalent to full health). Dichotomous valuations make the measure extremely sensitive to variation in the arbitrary threshold definition, which creates significant obstacles to cross-national comparisons and assessments of changes over time. Other health expectancies such as Years of Healthy Life, disability-adjusted life expectancy or health capital use polychotomous or, in principle, continuous valuations (see Figure 2). For those health expectancies that do not use arbitrary dichotomous schemes, the valuation approach can be distinguished further by five characteristics: (a) whose values are used, individuals in health states, the general public, healthcare providers or household members of individuals in health states; (b) what type of valuation question is used, such as the standard gamble, time trade-off, person trade-off, or visual analogue; (c) how are health states presented for the elicitation of valuations, that is, with what type of description and what level of detail, including some selection of domains; (d) what range of health states from mild to severe are valued at the same time; and (e) what combination of valuation questions are used and, more generally, what type of deliberative process is undertaken, if any. All of these questions continue to be debated extensively in the health economics literature and are not discussed in detail here.

Fifth, values other than health state valuations also may be incorporated explicitly into health expectancies. The calculation of health capital includes individuals' discount rates for future health (Cutler and Richardson 1997, 1998). In principle, other values might be included, such as age-weights, which allow years lived at different ages to take on different values, or equity weights, which allow years lived by one group or another take on different values. Examples of applications of these values are cited below in reference to health gaps. Incorporating other values into the design of a summary measure will usually require strong assumptions about the separability of population health across time and across persons. Such separability has been challenged in the literature on quantifying the benefits of health interventions – see for example, the debate on healthy year equivalents (HYEs) and quality adjusted life years (QALYs) (Mehrez and Gafni 1990, Johannesson et al. 1993, Cuyler and Wagstaff 1993).

Attributes of Health Gaps

Since Dempsey (1947), there has been extensive development of various measures of years of life lost due to premature mortality (e.g., Murray 1996, Romeder et al. 1977). Years of life lost measures are all measures of a mortality gap, the area between the survivorship function and some implied population norm for survivorship (area C in Figure 1). Health gaps extend the notion of mortality gaps to include time lived in health states worse than ideal health. A variety of health gaps have been proposed and measured (Murray 1996, World Bank 1993, Hyder et al. 1998, Ghana Health Assessment Project Team 1981) and many others could be derived. Health gaps can be distinguished on the basis of four dimensions.

First, what is the implied population health target or norm? Health gaps measure the difference between current conditions and a selected target. The explicit or implicit target is a critical characteristic of any health gap. Despite the obvious importance of choosing the health target, in some cases such as HeaLYs (Hyder et al 1998) the population target is not stated nor is it easily calculated. The original formulation of many mortality gaps was constructed in terms of the loss to each individual. The aggregate population implications of the loss due to premature mortality for each individual have been poorly appreciated sometimes. For example, Murray (1996) has shown that for many mortality gap measures and health gaps, the implied target may change as the mortality level changes, making direct comparisons between communities impossible, a problem discussed further below.

Second, as with health expectancies, a key issue is how health states worse than ideal health are defined and measured. In disability-adjusted life years, for example, health states are multi-dimensional and are based on both observations and self-perceptions of performance in different domains.

Third, health gaps can be distinguished by the method used to value time spent in different health states. Health state valuations are essential so that time lived in health states worse than ideal health can be compared with time lost due to premature mortality. Currently used gap measures are calculated using continuous valuations for health states, but the empirical methods used to elicit these valuations differ substantially.

Fourth, other values have commonly been included in gap measures; for example, age-weights and time preference are used in some variants of disability-adjusted life years (Murray and Lopez 1996, Barendregt et al. 1996, Murray and Acharya 1997). Health gaps that include explicit equity weights also have been proposed (Williams 1997).

Criteria for Evaluating Summary Measures

Given the extensive interest in summary measures and the range of health expectancies and health gaps, one way to proceed is to propose a minimal set of desirable properties that summary measures should have and evaluate available summary measures against these criteria. The minimalist set of desirable properties for summary measures is likely to vary depending on the intended use; that is, a summary measure most appropriate for comparisons of population health over time may not be most appropriate or even acceptable for reporting on the contribution of diseases, injuries and risk factors to population ill-health. In this section, we first address the question of choosing an appropriate summary measure for comparative purposes (uses 1 through 3 from above) and then address the question of choosing a summary measure that can be decomposed into the contributions of different diseases, injuries and risk factors (for uses 5 through 7). The purpose of this section is not to identify the best summary measure for each use, but rather to begin to define an explicit framework for making these choices.

There is a common-sense notion of population health such that, for some examples, everybody could agree that one population is healthier than another, or that the health of some particular population is getting worse or better. For instance, if two populations are identical in every way except that infant mortality is higher in one, then we expect that everybody would agree that the population with the lower infant mortality is healthier. On the basis of this type of common-sense notion, we can develop some very simple criteria for evaluating summary measures of health. As we

show below, however, even these simple criteria will lead us to some rather thought-provoking conclusions. Ultimately, we will need a more formal conceptual structure for constructing criteria to select summary measures.

Much of the discussion on the design of summary measures has been linked closely to the goal of maximizing gain in a summary measure in the face of a budget constraint. Inevitably, this has led to methods for constructing summary measures that emphasize the value choices involved in the allocation of scarce resources, for example the use of the person trade-off technique for measuring health state valuations (Nord 1992, Murray 1996). Many have rightly focused on a range of values that are relevant to allocating scarce resources that may enhance individuals' health (Nord et al. 1999). The problem is that many of these considerations bring us far from the common-sense statement that population A is healthier than B.

At least for the purposes of comparative statements on health, it may be necessary to distance the development of summary measures from the complex values that must be considered in the allocation of scarce resources. To put it in other words, we can quite reasonably choose to measure population health in one way and conclude that scarce resources should not be allocated simply to maximize population health as so measured – but surely population health as measured would at least be an argument in the social objective function for allocating scarce health resources. Although tentative, we believe that we can construct summary measures for comparative purposes based on an application of Rawls' principle of a veil of ignorance (Rawls 1971). In this construct, an individual behind a veil of ignorance does not know who he or she is in a population. We propose that the relation "is healthier than" can be defined such that population A is healthier than population B if and only if an individual behind a veil of ignorance would prefer to be one of the existing individuals in population A rather than an existing individual in population B, holding all non-health characteristics of the two populations to be the same.ⁱⁱ We emphasize that the principle of choice behind the veil of ignorance does not mean choosing to join one of the populations as an *additional* member, but rather choosing between two populations with the knowledge that you would be one of the *current* members of either population, but ignorance at the moment of choice as to which particular member you would be in either population. Implementing the veil of ignorance approach to selecting the criteria for a summary measure of population health would have many far-reaching implications. We explore only a few in this paper.

Based on the veil of ignorance argument and consonant with common-sense notions of population health, we argue that there are minimally five criteria that a summary measure should fulfil **for comparative purposes**.

Criterion 1. If age-specific mortality decreases in any age-group, *ceteris paribus*, then a summary measure should improve (i.e. a health gap should decrease and a health expectancy should increase). This criterion could be weakened to say that if age-specific mortality decreases in any age-group, *ceteris paribus*, then a summary measure should improve or stay the same. The weaker version would allow for deaths beyond some critical age to leave a summary measure unchanged. Measures such as potential years of life lost would then fulfil the weak form of the criterion.

By inspection, all health expectancies fulfil the strong form of this criterion but some health gaps do not. For example, to calculate the mortality component of HealLYs, the local life expectancy at each age is used to define the gap due to a death at each age. It can be demonstrated that HealLYs violate Criterion 1. For the purposes of

illustration, imagine two hypothetical populations with linear survivorship functions, as represented by the bold diagonals in Figure 3. For the population represented in the first diagram, life expectancy at birth (the area under the survivorship function) is 25 years, while the second population has a life expectancy at birth of 37.5 years. Based on the survivorship function, $s(x)$, we can compute, for each population, the life expectancy at each age, $e(x)$, namely, the area under the survivorship function to the right of age x divided by $s(x)$. The implied population norm, $G(\cdot)$, is defined such that $G(x+e(x)) = s(x)$. The population survivorship norm, $G(\cdot)$, is shown as the diagonal line to the right of the survivorship curve. In the population with the life expectancy of 37.5, which has a lower age-specific mortality rate at every age than the population with the life expectancy of 25, the health gap shown as the shaded area has actually increased.

Criterion 2. If age-specific prevalence of some health state worse than ideal health increases, *ceteris paribus*, a summary measure should get worse. Imagine two populations, A and B, with identical mortality, incidence and remission for all non-fatal health states, but with a higher prevalence of paraplegia in population A. Behind a veil of ignorance, an individual will prefer to be a member of population B; likewise our common-sense notion of population health would force us to conclude that B is healthier than A. Health expectancies and health gaps that are calculated using only incidence and remission rates for non-fatal health states will not fulfil this criterion, whereas prevalence-based health expectancies and health gaps will.

Criterion 3. If age-specific incidence of some health state worse than ideal health increases, *ceteris paribus*, a summary measure should get worse. Imagine two populations, A and B, with identical mortality, prevalence and remission, but with a higher incidence of blindness in A than in B. We must conclude that B is healthier than A. Incidence-based health expectancies and health gaps would fulfil this criterion.

Taking criteria 2 and 3 together, we are led to a very uncomfortable conclusion: no existing summary measure fulfils both criteria.ⁱⁱⁱ According to the conventional wisdom in health statistics, incidence-based measures are better for monitoring current trends, and they are more logically consistent for summary measures because mortality rates describe incident events. Prevalence measures are widely recognized as important for planning current curative and rehabilitative services, while incidence-based measures are more relevant to planning prevention activities. While these long-standing arguments have their merits, they are not answering the question posed here, which asks what it means for one population to be healthier at a moment in time than another. It seems undeniable that the common dichotomy between incidence-based or prevalence-based measures does not reflect the composite judgement that an individual behind a veil of ignorance would make on which population is healthier.

As one possible solution to this dilemma, we could estimate cohort health expectancy at each age x , which would reflect both incidence and prevalence. That is, the health expectancy of 50 year-olds will depend on both the current prevalence of conditions among individuals in this age group, as well as the current and future incidence, remission and mortality rates that this cohort will face. A period health expectancy at each age x which reflects incidence and prevalence could also be constructed to provide a measure based only on currently measurable aspects of health. A summary measure for the population could then be based on some aggregation of the cohort or period health expectancies at each age, such as a simple average across all individuals

in the population. This aggregate measure would reflect both incidence and prevalence and would be age-structure dependent. Although the mechanics of constructing this measure must be developed more fully, we believe that it is one potential solution to the problem posed jointly by criteria 2 and 3.

Criterion 4. If age-specific remission for some health state worse than ideal health increases, *ceteris paribus*, a summary measure should improve. The argument for this criterion is essentially identical to the argument for criterion 3.

Criterion 5. If the severity of a given health state worsens, *ceteris paribus*, then a summary measure should get worse. For some multi-dimensional constructions of health states, it could be argued that severity cannot change without the health state changing. In practice, most summary measures based on categorizing the population into a finite number of discrete health states worse than ideal health allow severity to change for individuals in a given state. This criterion is particularly important for assessing the performance of health systems where much of the health expenditure in high-income countries may be directed to interventions that reduce the severity of symptoms without changing mortality, incidence or remission rates. Disability-free life expectancy, impairment-free life expectancy and dementia-free life expectancy that use arbitrary dichotomous weights do not fulfil criterion 5.

Other Desirable Properties for Summary Measures

Summary measures for comparative purposes are meant to inform many policy discussions and debates. The intended widespread use of summary measures leads to several desirable attributes in addition to the basic criteria described above. These are not attributes based on formal or informal arguments about whether a population is healthier than another but rather on practical considerations. Unfortunately, it is possible that no measure can have all the desirable attributes.

1) Summary measures should be comprehensible and feasible to calculate for many populations. It is of little value developing summary measures that will not be used to inform the health policy process. The nearly universal use of a very complex abstract measure, namely period life expectancy at birth, demonstrates that comprehensibility and complexity are different. The interest of the popular press in **DALE (Washington Post editorial, emma can you search lexis-nexus)**, probably because health expectancies build on life expectancy, is one indication of the comprehensibility of health expectancies. Health gaps are perhaps less familiar to many but the concept is relatively simple and communicable.

2) It would be convenient if summary measures were linear aggregates of the summary measures calculated for any arbitrary partitioning of sub-groups. Many decision-makers, and very often the public, desire information that is characterized by this type of additive decomposition. In other words, they would like to be able to answer what fraction of the summary measure is related to health events in the poor, in the uninsured, in the elderly, in children, and so on. Additive decomposition is also often appealing for cause attribution, which is discussed in detail below. Additive decomposition can be achieved for health gaps but not for health expectancies. For example, we can report the number of DALYs in a population for those ages 0 to 4 years, and those ages 5 years and older, and the sum of these two numbers will equal the total health gap in the population; it is unclear, on the other hand, how to combine the disability-adjusted life expectancy for everybody ages 0 to 4 with the disability-adjusted life expectancy for everybody ages 5 years and older into a meaningful

number. Techniques for estimating the contribution of changes in age-specific mortality rates to a change in life expectancy have been developed (e.g. Arriaga 1984) but they do not have the property of additive decomposition.

Calculating the Contribution of Diseases, Injuries and Risk Factors to Summary Measures of Population Health

Another fundamental goal in constructing summary measures and one that may, in fact, explain the increasing attention to summary measures is to identify the relative magnitude of different health problems, including diseases, injuries and risk factors – which corresponds to uses 5 through 7 above. There are two dominant traditions in widespread use for causal attribution: categorical attribution and counterfactual analysis. There has been remarkably little debate or discussion on the advantages and disadvantages of the two approaches or the inconsistency of using both approaches in the same analysis. An example of using both approaches in the same body of work is the Global Burden of Disease Study (Murray and Lopez 1996a). Burden attributable to diseases and injuries has been estimated using categorical attribution whereas burden attributable to risk factors or diseases such as diabetes that act as risk factors has been estimated using counterfactual analyses.

1) Categorical attribution

An event such as death or the onset of a particular health state can be attributed categorically to one single cause according to a defined set of rules. In cause-of-death tabulations, for example, each death is assigned to a unique cause according to the rules of the International Classification of Disease, even in cases of multi-causal events. For example, in ICD-10, deaths from tuberculosis in HIV-positive individuals are assigned to HIV. This categorical method of representing causes is the standard method used in published studies of health gaps such as the Global Burden of Disease (Murray and Lopez 1997a). The advantage of categorical attribution is that it is simple, widely understood and has the property of additive decomposition. The disadvantage is well illustrated by multi-causal events such as a myocardial infarction in a diabetic, or liver cancer resulting from chronic hepatitis B.

For categorical attribution to work a classification system is required. A classification system has two key components: a set of mutually exclusive and collectively exhaustive categories and a set of rules to assign events to these categories. For diseases and injuries, the International Classification of Disease has been developed and refined over the course of nearly 100 years. No classification system for other types of causes such as physiological, proximal or distal risk factors has been developed thus far.

2) Counterfactual analysis

The contribution of a disease, injury or risk factor can be estimated by comparing the current level and future levels of a summary measure of population health with the levels that would be expected under some alternative hypothetical scenario. This scenario might be a counterfactual distribution of risk or the extent of a disease or injury. The models used in counterfactual analyses may be extremely simple or, in the case of some risk factors, which can have complex time and distributional characteristics, the models can be quite complex. The validity of the estimate depends on the validity of the model used to predict the counterfactual scenarios.

Many types of counterfactuals may be used for this type of assessment:

- a) The effect of small changes in the disease, injury or risk factor can be assessed and the results expressed as the elasticity of the summary measure with respect to changes in the disease, injury or risk factor, or as a numerical approximation of the partial derivative of the summary measure (Hill et al. 1996, Mathers 1999). For health expectancies, the effect of a small change can be assessed in terms of a change in the period health expectancy, and for health gaps the change can be evaluated in terms of the change in the present value of future health gaps.
- b) The change in a summary measure expected with complete elimination of a disease or injury is another form of counterfactual analysis. A number of studies have presented results on cause-deleted health expectancies (Colvez and Blanchet 1983, Mathers 1992, 1997, 1999, Nusselder 1998, Nusselder et al. 1996). Wolfson (1996) has calculated attribute-deleted health expectancies (deleting types of disabilities rather than causes). For selected risk factors such as tobacco or alcohol use, counterfactual analyses based on the complete elimination of the risk are possible (Murray and Lopez 1999), but for other risk factors such as blood pressure, complete elimination is not meaningful.
- c) More generally, Murray and Lopez (1999) have developed a classification of various counterfactual risk distributions that can be used for these purposes, including the theoretical minimum risk, the plausible minimum risk, the feasible minimum risk and the cost-effective minimum risk. The examples of tobacco and alcohol have been used to explore the implications of using these different types of counterfactual distributions to define attributable burden and avoidable burden.
- d) Another form of counterfactual analysis is intervention analysis where the change in a summary measure from the application of an intervention is estimated.

Counterfactual analysis of summary measures has a wide spectrum of uses from the assessment of specific policies or actions to more general assessments of the contribution of diseases, injuries or risk factors. We have described four different types of counterfactual analyses, but an important consideration that is independent of the type of counterfactual used is the duration of the counterfactual. For example for tobacco use, the counterfactual could be one year in which the population does not smoke, followed by a return to the status quo at the end of this year. The health effects that are due to one year of tobacco consumption would then be traced out in terms of changes in future health expectancies or future health gaps. Alternatively, the counterfactual change could be longer, such as a permanent change to a state of no tobacco consumption. In the case of a permanent shift to the counterfactual, there could be an infinite stream of future changes in health expectancies or health gaps. If the summary measure of population health does not include discounting of future health, then the change in a summary measure attributable to a permanent change in a risk factor or a disease could be infinite. For this reason alone, the convention of analyzing a single year counterfactual may be preferable.

For analyzing the contribution of diseases, injuries or risk factors, three options as illustrated by Table 1 are possible. Population health can be summarized using health expectancies and health gaps, and cause attribution for diseases and injuries can be assessed using categorical attribution or counterfactuals. Because there is no

classification system for risk factors, they can only be assessed using the counterfactual approach. Even for diseases and injuries, it is not possible to use categorical attribution with a health expectancy. Positive health cannot be categorically assigned to diseases or injuries. For the three options in Table 1, what are the advantages and disadvantages?

For causal attribution, a very high premium is put on additive decomposition by most users. Additive decomposition means that the total level of the summary measure equals the contribution of a set of mutually exclusive and collectively exhaustive causes such as the ICD list of diseases and injuries. If additive decomposition is a critical property, then the contribution of diseases and injuries can only be assessed using health gaps.

The counterfactual method for calculating the contribution of diseases, injuries and risk factors has many advantages. It is conceptually clearer, it solves problems of multi-causality and it is consistent with the approach for evaluating the benefits of health interventions (Murray and Lopez 1999). While the counterfactual approach has these advantages both for health gaps and health expectancies, health expectancies do not give results that are related to the absolute magnitude of gain in healthy life. For health gaps, on the other hand, the counterfactual approach produce results that reflect the absolute age and sex distribution of the population. For causal attribution to be used to inform debates on R&D priorities, selection of national health priorities for action, and health curriculum development, it can be argued that a method of causal attribution should give an ordinal ranking of causes that is identical to the ordinal ranking of absolute years of healthy life gained by a population through cause elimination (or appropriate counterfactual change for a risk factor). In other words, for cause decomposition intended to inform public health prioritization, the absolute numbers attributable to a cause matter.

With a health gap, the natural denominator for categorical attribution is the current health gap. But the denominator for counterfactual analysis of causal attribution would more logically be the present value of the total health gap from all causes for the currently living population. Further consideration is required to develop appropriate denominators for counterfactual assessments of risk factors for environmental changes that may affect many generations.

Discussion

In this paper, we have put forward a basic framework for characterizing and evaluating different types of summary measures. In choosing summary measures for a range of different uses, it is critical to understand the important differences between the various types of available summary measures, but also to distinguish clearly between the range of summary measures themselves and the different types of instruments and data that may be used as inputs to estimating these summary measures. We have defined in this paper a skeletal set of five basic criteria that may be used as a starting point in evaluating summary measures. We hope that this set of criteria will evoke further debate on other possible criteria that may be useful to analysts and policy makers in choosing summary measures for policy applications.

Nevertheless, it is worth noting for comparative use that one health gap, HeaLYs does not even fulfil the most basic criterion that as mortality gets worse the measure gets worse. According to criterion 5, we should also reject measures with arbitrary dichotomous weights for time spent in health states less than full health including

disability-free life expectancy (DFLE), active life expectancy (ALE) and dementia-free life expectancy. A number of remaining health expectancies and health gaps fulfil four out of the five criteria, but no measure fulfils the prevalence and incidence criteria at the same time. For comparative uses, we may need to develop a new class of measure that reflects both prevalence and incidence such as the average age-specific health expectancy described above. It is very important to recognize that for other uses of summary measures, different criteria may be formulated with different implications for the design of summary measures.

Causal attribution is a key aspect of summary measures for several important uses outlined in this paper. For diseases and injuries, the development this century of the International Classification of Diseases allows a choice between categorical attribution and counterfactual analysis. The desirability of additive decomposition strongly favours the use of categorical attribution. On the other hand, the magnitudes from counterfactual analysis have a more direct and theoretically cogent interpretation. We suggest that in practice the only solution to this tension is to routinely report both categorical attribution and counterfactual analyses for diseases and injuries. All issues of multi-causal death, such as for diabetes mellitus, would be well captured in the counterfactual analysis even if the categorical attribution tends to ‘underestimate’ the problem. For risk factors – whether physiological, proximal or distal – no classification system presently exists, so that the only option is counterfactual analysis. There are many options for defining counterfactuals, and substantial work is needed to understand more fully the implications of adopting different approaches.

Improving the actual estimation of summary measures of population health depends not only on designing the most appropriate summary measure for a particular purpose, but also on improving the empirical basis for the epidemiology of fatal and non-fatal health outcomes including, attribution by cause, and improving the empirical basis for health state valuations. One critical hurdle for progress on the epidemiology of non-fatal health outcomes to construct summary measures for comparative purposes is an improved understanding of the determinants of the difference between self-reported and presumably self-assessed performance in selected domains of health and observed performance.

In proposing this framework for choosing summary measures, we have invoked both a common-sense notion whereby in some cases everybody could agree that one population was healthier than another, as well as a more formal mechanism for defining this choice, using Rawls’ notion of choice behind a veil of ignorance. There are some potentially important implications of the veil of ignorance framework for choosing a summary measure of population health for comparative purposes. For example, the current methods used to measure preferences for time spent in health states may not be entirely consistent with this framework, and modified methods would perhaps need to be developed. Clearly, it will be helpful to provide a more rigorous formal treatment of this approach.

As work on summary measures gathers speed, the uses and complexities of these measures are becoming more widely appreciated. In this paper, we have suggested that the application of simple criteria may lead us to reject some measures and develop new ones. An extensive developmental agenda exists; nevertheless, we believe that real world use of summary should not wait until all methodological issues have been resolved. For comparative purposes in the interim, we recommend the use

of a prevalence-based health expectancy using polychotomous or continuous weights for health states defined in terms of multiple domains of health – for example, disability-adjusted life expectancy. Given the present state of development of self-assessed health instruments and the difficulty of comparing across individuals and groups, we do not believe that such a health expectancy should be estimated on the basis of survey data from a single instrument. The challenge for estimating health expectancies is to integrate epidemiological information from a variety of sources to develop the most comparable and valid assessment. For purposes that require causal attribution, we believe that a health gap is strongly preferable. In the interim, the best options are to use DALYs[0,0] – DALYs calculated using a zero discount rate and uniform age weights – or DALYs[0.03,1] – calculated with a 3% discount rate and non-uniform age weights.

Literature Cited

Anand S, Hanson K. DALYs: Efficiency versus Equity. *World Development*, 1998 26 (2): 307-310.

Barendregt JJ, Bonneux L. Changes in incidence and survival of cardiovascular disease and their impact on disease prevalence and health expectancy. In: Mathers C., McCallum J, Robine J-M. (eds). *Advances in health expectancies: Proceedings of the 7th Meeting of the International Network on Health Expectancy (REVES), Canberra, February 1994*. Australian Institute of Health and Welfare: AGPS, Canberra: 1994, 345-354.

Barendregt JJ, Bonneux L, Van der Maas PJ. Health expectancy: an indicator for change? *Epidemiology of Human Health*, 1994, 48:482-7.

Bobadilla JL. *Searching for essential health services in low- and middle-income countries: A Review of Recent Studies on Health Priorities*. Human Development Department, World Bank, 1996.

Bone MR. International efforts to measure health expectancy. *Journal of Epidemiology and Community Health*, 1992, 46, 555-558.

Bowie C, Beck S, Bevan G, Raftery J, Silvertion F, and Stevens J. Estimating the burden of disease in an English region. *Journal of Public Health Medicine*, 1997, 19:87-92.

Branch LG, Guralnik JM, Foley DJ, Kohout FJ, Wetle TT, Ostfeld A, Katz S. Active life expectancy for 10,000 Caucasian men and women in three communities. *Journal of Gerontology*, 1991, 46(4): M145-50.

Bronnum HH. Trends in health expectancy in Denmark, 1987-1994. *Dan. Med. Bull.* 1998, 45, 217-221.

Brundtland, GH. The global burden of disease. Speech delivered on December 15, 1998, Geneva, 1998.

Chiang CL. *An index of health: mathematical models*. *Public Health Services Publications 1000 Series2. No.5*. Washington DC, National Center for Health Statistics, 1965.

- Colvez A, Blanchet M. Potential gains in life expectancy free of disability: a tool for health planning. *International Journal of Epidemiology*, 1983, 12: 224-229.
- Concha M, Aguilera, X, Albala C, et al. Estudio ecarga de enfermedad informe final. Estudio Prioridades de Inversio en salud Minsterio de Salud, 1996.
- Crimmins, EM., Saito, Y, Ingengneri, D. Trends in disability-free life expectancy in the United States, 1970-90. *Population and development review*, 1997, 23(3): 555-572.
- Cutler DM, Richardson E. The value of health. *American Economic Review*, 1998, 88(2): 97-100.
- Cutler DM, Richardson E. Measuring the health of the U.S. population. *Brookings Paper: Microeconomics*, 1997, 217-271.
- Cuyler AJ, Wagstaff A. QALYs versus HYE. *Journal of health economics* 1993, 311-323.
- Daniels, op.cit: Brock DW. Justice, Health Care, and the elderly. *Philosophy and Public Affairs*, 1989, 18(3): 297-312.
- Deeg DJH, Kriegsman D.M.W., van Zonneveld R.J. Trends in fatal chronic diseases and disability in the Netherlands 1956-1993. In: Mathers C., McCallum J., Robine J.-M. (eds). *Advances in health expectancies: Proceedings of the 7th Meeting of the International Network on Health Expectancy (REVES), Canberra, February 1994*. Australian Institute of Health and Welfare: AGPS, Canberra, 1994, 80-95.
- Dempsey M. Decline in tuberculosis: the death rate fails to tell the entire story. *American review of tuberculosis*, 1947, 56: 157-164.
- Erickson P, Wilson R, Shannon E. Years of healthy life. Hyattsville, Maryland: US National Center for Health Statistics, 1995.
- Erickson P. Evaluation of a population-based measure of quality of life: the Health and Activity Limitation Index (HALex). *Quality of Life Research*, 1998, 7:101-114.
- Fanshel S, Bush JW. A health-status index and its application to health services outcomes. *Operations research*, 1970, 18(6):1021-1066.
- Field M.J, Gold GM, eds. *Summarizing Population Health: Directions for the Development and Application of Population Metrics*. Institute of Medicine, Washington, D.C. National Academy Press, 1998.
- Fries J.F. Aging, natural death and the compression of morbidity. *New England Journal of Medicine*, 1980, 303:130-135.
- Fundación Mexicana Para La Salud. *Health and the Economy: Proposals for Progress in the Mexican Health System*. Mexico, Salud, 1995.
- Ghana Health Assessment Project Team. A quantitative method of assessing the health impact of different diseases in less developed countries. *International journal of epidemiology*, 1981, 10(1):72-80.
- Hill GB, Forbes WF, Wilkins R. The entropy of health and disease: dementia in Canada. 9th meeting of the International Network on Health Expectancy (REVES), Rome, December 11-13, 1996.
- Hyder AA, Rotllanat G, Morrow RH. Measuring the Burden of Disease: Healthy Life-Years. *American Journal of Public Health*, 1998, 88(2):196-202.

- Jamison DT. Forward to the *Global Burden of Disease and Injury Series*. In: Murray CJL, and Lopez A, eds. *The Global Burden of Disease*. Cambridge, Harvard University Press, 1996.
- Johannesson-M; Pliskin-JS; Weinstein-MC. Are healthy-years equivalents an improvement over quality-adjusted life years? *Med-Decis-Making*. 1993 Oct-Dec; 13(4): 281-6
- Katz S, Akpom CA, Papsidero JA, Weiss ST. Measuring the health status of populations. In: Berg RL, ed. *Health Status indices*. Chicago, Hospital Research and Educational Trust, 1973, 39-52.
- Katz S, Branch LG, Branson MH., Papsidero JA, Beck JC, Greer DS. Active life expectancy. *The New England Journal of Medicine*, 1983, 309(20):1218-24.
- Lozano R, Frenk J, Gonzalez MA. El peso de la enfermedad en adultos mayores, Mexico 1994. *Salud Publica de Mexico*, 1994, 38: 419-429.
- Lozano R, Murray CJL, Frenk J, and Bobadilla J. Burden of Disease Assessment and Health System Reform: Results of a study in Mexico. *Journal for International Development*, 1995, 7(3): 555-564.
- Mackenbach JP, Kunst AE, Lautenbach H, Oei YB, Bijlsma F. Gains in life expectancy after elimination of major causes of death: revised estimates taking into account the effect of competing causes. *Journal of Epidemiology and Community Health*, 1999, 53: 32-37.
- Mathers CD. Gains in health expectancy from the elimination of diseases among older people. *Disability and Rehabilitation*, 1999, (in press).
- Mathers C.D. Gains in health expectancy from the elimination of disease: a useful measure of the burden of disease? Paper presented to the 10th Meeting of the International Network on Health Expectancy (REVES), Tokyo, 9-11 October 1997.
- Mathers C.D. 1992. Estimating gains in health expectancy due to elimination of specified diseases. Fifth meeting of the International Network on Health Expectancy (REVES-5), Statistics Canada, Ottawa, 19-21 February 1992.
- Mathers CD, Robine J-M. How good is Sullivan's method for monitoring changes in population health expectancies. *Journal of Epidemiology and Community Health*, 1997a, 51: 80-86.
- Mathers CD, Robine J-M. How good is Sullivan's method for monitoring changes in population health expectancies? Reply. *Journal of Epidemiology and Community Health*, 1997b, 51:578-579.
- Mathers C, McCallum J, Robine J-M., eds. Advances in health expectancies: proceedings of the 7th meeting of the International Network on Health Expectancy (REVES), Canberra. Canberra, Australian Institute of Welfare: AGPS, 1994.
- Mehrez A, Gafni A. Quality-adjusted life-years, utility theory and healthy-year equivalents. *Medical Decision Making* 1989. 9:142-149.
- Mooney G, Irwig L, Leeder S. Priority setting in health care: unburdening from the burden of disease. *Aust NZ J Public Health* 1997, 21(7):680-681.
- Murray CJL. Rethinking DALYs. In: Murray C.J.L. and Lopez A. eds. *The Global Burden of Disease*. Cambridge, Harvard University Press, 1996.

- Murray CJL. and Acharya A.K. Understanding DALYs. *Journal of Health Economics*, 1997, 16: 703-730.
- Murray CJL, and Lopez A. On the comparable quantification of health risks: lessons from the Global Burden of Disease Study. *Epidemiology*, 1999 (in press).
- Murray CJL, and Lopez A. Mortality by cause for eight regions of the world: Global Burden of Disease Study. *Lancet*, 1997a, 349: 1269-1276.
- Murray CJL, and Lopez A. Regional patterns of disability-free life expectancy and disability-adjusted life expectancy: Global Burden of Disease Study. *Lancet*, 1997b, 349: 1347-1352.
- Murray CJL, and Lopez A. Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. *Lancet*, 1997c, 349: 1436-1442.
- Murray CJL, and Lopez A. Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. *Lancet*, 1997d, 349: 1498-1504.
- Murray CJL, and Lopez A. *The Global Burden of Disease*. Cambridge, Harvard University Press, 1996.
- Murray CJL, and Lopez A. *Global Health Statistics*. Cambridge, Harvard University Press, 1996.
- Murray CJL, and Lopez A. Evidence-based health policy—lessons from the Global Burden of Disease Study. *Science*, 1996, 274:740-743.
- Mutafova M, van-de-Water HP, Perenboom RJ, Boshuizen HC, and Maleshkov C, Health expectancy calculations: a novel approach to studying population health in Bulgaria. *Bulletin of the World Health Organization*, 1997.
- Nord E. The person trade-off approach to valuing health care programs. *Medical decision making*, 1995, 15(3): 201-208.
- Nord E, Pinto JL, Richardson J, Menzel P, Ubel P. Incorporating societal concerns for fairness in numerical valuations of health programmes. *Health Economics*, 1999, 8: xx-xx.
- Nusselder WJ. Compression or expansion of morbidity? A life-table approach. *Ph.D. Thesis*. Erasmus University, Rotterdam, 1998.
- Nusselder WJ, van der Velden K, Sonsbeek JLA, Lenior ME, van den Bos GAM. The elimination of selected chronic diseases in a population: the compression and expansion of morbidity. *American Journal of Public Health*, 1996, 86(2); 187-193.
- Piot M, Sundaresan TK. *A linear programme decision model for tuberculosis control. Progress report on the first test-runs*. Geneva, World health Organization, 1967, WHO/TB/Tech.Information/67.55.
- Preston SH. Health indices as a guide to health sector planning: a demographic critique. In Gribble JN, Preston SH, eds. *The epidemiological transition. Policy and planning implications for developing countries*. Washington D.C., National Academic Press, 1993.
- Rawls J. *A theory of justice*. Cambridge: Harvard University Press, 1971.
- República de Colombia Ministerio de Salud. *La Carga de la Enfermedad en Columbia*. Bogota, Ministerio de Salud, 1994.

- Ritchie K, Jagger C, Brayne C, Letenneur L. Dementia-free life expectancy: preliminary calculations for France and the United Kingdom. In Robine J-M, Mathers CD, Bone MR, Romieu I, eds. *Calculation of health expectancies: harmonization, consensus achieved and future perspectives*. London, John Libbey Eurotext, 1993.
- Robine JM, et al., eds. *Calculation of health expectancies: harmonization, consensus achieved and future perspectives*. London, John Libbey Eurotext, 1993.
- Robine, J.M., Blanchet, M., Dowd, EJ, eds. *Health Expectancy*. OPCS, HMSO, London, 1992.
- Robine JM, Michel JP, Branch LG. Measurement and Utilization of Healthy Life Expectancy: conceptual issues. *Bulletin of the World Health Organization*, 1992, 70(6): 791-800.
- Rogers A, Rogers RG, Branch LG. A Multistate Analysis of Active Life Expectancy. *Public Health Reports*, 1989, 104(3):222-226.
- Romeder JM, McWhinnie JR. Potential years of life lost between ages 1 and 70: an indicator of premature mortality for health planning. *International journal of epidemiology*, 1977, 6:143-151.
- Ruwaard D. and Kramers PGN. *Public health status and forecasts*. The Hague: National Institute of Public Health and Environmental Protection, 1998.
- Sanders BS. Measuring community health levels. *American Journal of Public Health*, 1964, 54:1063-70.
- Shryock HS, Siegel JS. *The methods and materials of demography: the life table*. Washington D.C., U.S. Bureau of the Census, October 1971, 443-6.
- Sihvonen AP, Kunst AE, Lahelma E, Valkonen T, and Mackenbach JP. Socioeconomic inequalities in health expectancy in Finland and Norway in the late 1980s. *Soc.Sci.Med.* 1998, 47, 303-315.;
- Sullivan DF. A single index of mortality and morbidity *HSMHA health reports*, 1971, 86:347-354.
- Sullivan DF. *Conceptual problems in developing an index of health*. US public Health Service Publication Series No. 1000. *Vital and Health Statistics Series 2. No. 17*. National Center for Health Statistics, 1966.
- Valkonen T. Sihvonen AP. and Lahelma E. Health expectancy by level of education in Finland. *Soc.Sci.Med.* 1997, 44, 801-808.
- Wilkins R, Adams OB. Quality-adjusted life expectancy: weighting of expected years in each state of health. In Robine J-M, Blanchet M, Dowd JE (eds) *Health expectancy*, OPCS studies on medical and population subjects no. 54, HMSO, London, 1992.
- Williams A. Intergenerational equity. An exploration of the 'Fair Innings' Argument. *Health Economics*, 1997 6(2): 117-32.
- Williams A. Calculating the global burden of disease: time for a strategic reappraisal? *Health Economics*, 1999, 8: 1-8.
- Wolfbein S. The length of Working Life. *Population Studies*, 1949, 3:286-94.
- Wolfson M.C. Health-adjusted life expectancy. *Health reports*, 1996, 8(1): 41-45.

World Bank. *World Development Report 1993: investing in health*. New York: Oxford University Press for the World Bank, 1993.

World Health Organization. *Investing in health research and development. Report of the Ad Hoc Committee on Health Research Relating to Future Intervention Options*. Geneva: World Health Organization, 1996.

ⁱ Figure 1 is only a correct representation of a health gap in absolute terms in a stable population with zero growth.

ⁱⁱ Formally, the assumption that the relation “is healthier than,” conditional on fixed levels of non-health characteristics, does not vary depending on what these fixed levels are requires separability of the health-related characteristics of a population from non-health related characteristics. Questions of separability are indeed central to a number of conceptual issues underlying the construction of summary measures, but these questions fall beyond the scope of this paper.

ⁱⁱⁱ Health capital at age x , proposed as a measure of a cohort’s health by Cutler and Richardson (1997), is a subjective discounted cohort health expectancy at age x . While it has not been proposed as a summary measure of population health, it includes both prevalence and subjective expected incidence in its arguments.

Table 1. Approaches to analysing the contributions of diseases, injuries or risk factors to summary measures of population health.

	Categorical attribution	Counterfactual analysis
Health expectancies		<ul style="list-style-type: none"> • Diseases • Injuries • Risk Factors
Health gaps	<ul style="list-style-type: none"> • Diseases • Injuries 	<ul style="list-style-type: none"> • Diseases • Injuries • Risk Factors

Figure 1. The survivorship curve.

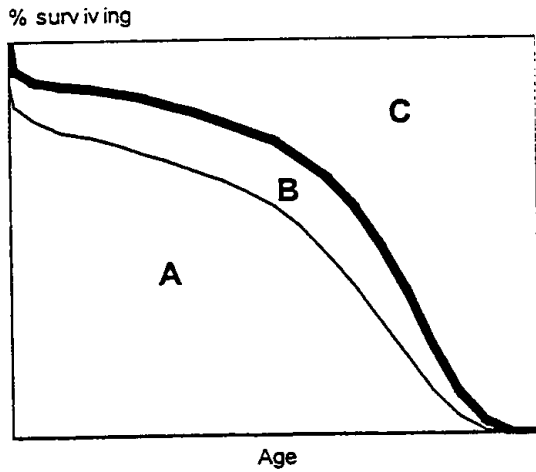


Figure 2. Valuations of time spent in health states worse than perfect health. The dotted line indicates dichotomous valuations, as in disability-free life expectancy, while the solid line indicates categorical valuations, as in disability-adjusted life expectancy.

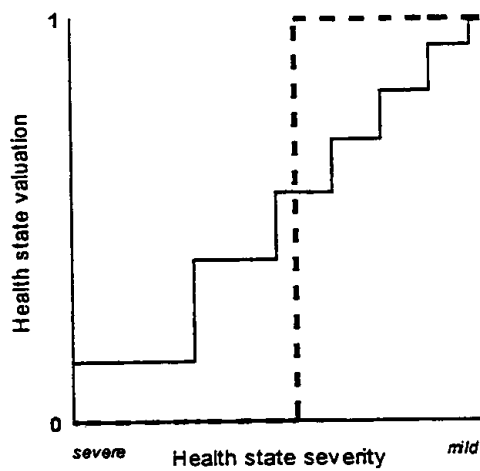


Figure 3. Survivorship curves in two hypothetical populations. The first graph indicates a population with life expectancy at birth of 25 years, while the second graph indicates a population with life expectancy of 37.5 years. In each graph, the bold diagonal represents a survivorship curve. The diagonal to the right of the survivorship curve indicates a survivorship norm in each population that is based on actual life expectancy at each age.

